Genomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Original Article

The first draft genome of *Lophophorus*: A step forward for Phasianidae genomic diversity and conservation

Kai Cui[a], Wujiao Li[b], Jake George James[a], Changjun Peng[a], Jiazheng Jin[b], Chaochao Yan[b], Zhenxin Fan[a], Lianming Du[c], Megan Price[b], Yongjie Wu[a,*], Bisong Yue[a,*]

[a] Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, Sichuan, China
[b] Sichuan Key Laboratory of Conservation Biology on Endangered Wildlife, College of Life Sciences, Sichuan University, Chengdu 610065, Sichuan, China
[c] Institute for Advanced Study, Chengdu University, Chengdu, Sichuan 610106, China

ARTICLE INFO

Keywords:
Chinese monal
Lophophorus lhuysii
Genome
De novo assembly
Adaptation

ABSTRACT

The monal genus (*Lophophorus*) is a branch of Phasianidae and its species inhabit the high-altitude mountains of the Qinghai-Tibet Plateau. The Chinese monal, *L. lhuysii*, is a threatened endemic bird of China that possesses high-altitude adaptability, diversity of plumage color and potentially low reproductive life history. This is the first study to describe the monal genome using next generation sequencing technology. The Chinese monal genome size is 1.01 Gb, with 16,940 protein-coding genes. Gene annotation yielded 100.93 Mb (9.97%) repeat elements, 785 ncRNA, 5,465,549 bp (0.54%) SSR and 15,550 (92%) genes in public databases. Compared to other birds and mammals, the genome evolution analysis showed numerous expanded gene families and positive selected genes involved in high-altitude adaptation, especially related to the adaptation of low temperature and hypoxia. Consequently, this gene data can be used to investigate the molecular evolution of high-altitude adaptation in future bird research. Our first published genome of the genus *Lophophorus* will be integral for the study of monal population genetic diversity and conservation, genomic evolution and Galliformes species differentiation in the Qinghai-Tibetan Plateau.

1. Introduction

The Chinese monal (*Lophophorus lhuysii*,Fig. 1) chiefly inhabits harsh highland meadows from 3500 to 4500 m elevation in the Hengduan Mountains (eastern Qinghai-Tibetan Plateau), south-west China. It has a narrow distribution and is restricted to eastern Tibet, south-east Qinghai, western Sichuan, northern Yunnan and southern Gansu (provinces) [1–4]. It was estimated that the population density was between 1.32 and 1.58/km^2 30 years ago [5]. Unfortunately, there has been no recent evidence of its presence in Tibet or Qinghai Provinces. Given the small population and ongoing decline, the species is considered as Vulnerable by the International Union for the Conservation of Nature Red List of Threatened Species [6]. It is also listed on Appendices I of CITES [7] and is a nationally protected Class I species in China.

The Chinese monal has two close relatives, the Himalayan monal (*L. impejanus*) and the Sclater's monal (*L. sclateri*), which together form the world's extant genus of monal (*Lophophorus)*. Zhan et al. [8] studied *Lophophorus* evolution by comparing its molecular phylogeny, distribution patterns and morphology, and have suggested that the genus originated in the Hengduan Mountains and broader Qinghai-Tibetan Plateau, which are hotspots for Galliformes species differentiation. Geographic barriers, topographic features, typically low temperatures, low oxygen and strong ultraviolet radiation throughout the Qinghai-Tibetan Plateau has influenced species differentiation and high-altitude adaptability of the monals and its closely related genera, such as *Tragopan*, *Pucrasia* and *Ithaginis* [9, 10].

This study aimed to describe the first monal genome and to provide important molecular material for the study of the monal genus and Phasianidae family. We conducted phylogenetic and comparative genomic analyses of the Chinese monal with other described bird species' genomes to increase our knowledge of the Chinese monal's high-altitude adaptability and evolutionary history. These analyses can provide valuable information for future studies of the different molecular mechanisms in bird high-altitude adaptation in the Qinghai-Tibetan Plateau. In particular, the results and material provided by our study can assist in researching the genomic evolution and genetic diversity of *Lophophorus* in the Hengduan Mountains to conserve and monitor monal populations and species viability.

* Corresponding authors.
E-mail addresses: wuyongjie@scu.edu.cn (Y. Wu), bsyue@scu.edu.cn (B. Yue).

https://doi.org/10.1016/j.ygeno.2018.07.016
Received 15 March 2018; Received in revised form 22 July 2018; Accepted 25 July 2018
0888-7543/ © 2018 Elsevier Inc. All rights reserved.

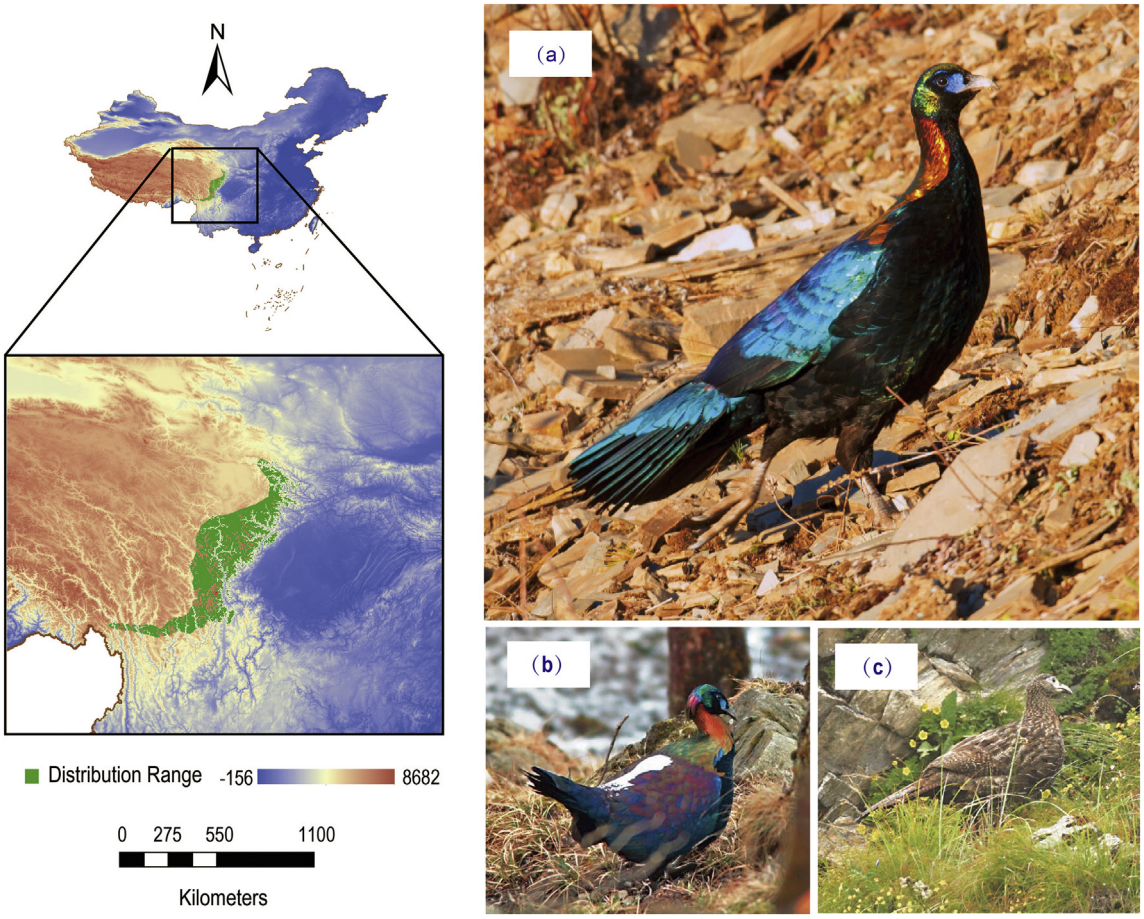

Fig. 1. The latest distribution map of the Chinese monal in China based on specimen and field records in China.
§ (a) And (b) A male with brightly colored feathers. (a) A female with brown feathers. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2. Materials and methods

2.1. Sample collection, sequencing and assembly

Blood samples were collected from a healthy, wild male Chinese monal (*L. lhuysii*: NCBI Taxonomy ID: 228248) from the Fengtongzhai National Nature Reserve, Sichuan Province, China. The collected blood was used for Genomic DNA extraction, isolation and sequencing. We used a whole genome shotgun approach on the Illumina HiSeq™ 2000 platform to sequence the genome. We constructed two paired-end libraries with insert sizes of about 230 base pairs (bp) and 500 bp, and three mate-paired libraries with insert sizes of 2 kb, 5 kb and 10 kb. After quality control and filtering, we obtained 174.66 Gb (134.34 ×) of clean data (Table 1).

Before assembly, a 17-Kmer analysis [11] was performed using Kmerfreq for estimating the genomic size of the Chinese monal with a 500 bp short-insert library. The frequency of each K-mer was calculated based on the clean reads. K-mer depth distribution was plotted against the sequence depth gradient and followed a Poisson distribution. Finally, genomic size was estimated by using the following formula: G = K_num/K_depth; where K_num is the total number of K-mer and K_depth is the frequency that occurs more than the others.

The *de novo* genome assembly used the ALLPATHS-LG pipeline [12], which optimizes parameters and employs the de Brujin graph. The main steps of the pipeline are as follows: First, base errors were identified and corrected using the frequencies and quality scores in reads; next, ALLPATHS built unipaths and estimated the copy number of unipaths using read depth to choose the seed unipath; then we built neighborhood paths to conduct neighborhood assembly; later these paths were glued together to form a global assembly graph; lastly, contigs were extracted from this global assembly and built into scaffolds to complete final assembly.

Table 1
Statistics of genome sequencing of the Chinese monal.

	Insert size	Total data (G)	Read length	Sequence coverage (X)	Average Q20 (%)
Pair-end libraries	230 bp	64.34	150	49.49	96
	500 bp	30.95	150	23.81	91
		2.38	150	1.83	89
Illumina reads	2 kb	31.65	150	24.34	95
	5 kb	23.08	150	17.75	95
	10 kb	22.26	150	17.12	95
Total	–	174.66	–	134.34	–

2.2. Estimation of genome completeness and assembly quality

In order to evaluate the quality of the assembly, the Chinese monal blood transcriptome was sequenced using Illumina HiSeq™ 2000 and was *de novo* assembled in Trinity (v2.1.1) [13]. The assembled transcripts were aligned to the assembly sequences using BLAT (v36) with default parameters (except an identity cutoff of 80%). Additionally, CEGMA (v2.5) [14] and BUSCO (v1.22) [15] were used to evaluate both the assembly quality and the genome completeness.

ARTICLE IN PRESS

2.3. Annotation

2.3.1. Repeat elements annotation

Transposable elements (TEs) and other repetitive elements were predicted in the Chinese monal genome by combining homology-based and *de novo* methods in RepeatMasker (v4.0.5) [16] and RepeatModeler (v1.0.8). In the homology-based approach, RepeatMasker was used to identify and classify repeat elements against the RepBase library and extract the information about homologous repeat elements. Then, RepeatMasker employed a *de novo* prediction against a *de novo* repeat sequences database constructed by RepeatModeler. By combining the data obtained from the two annotation approaches using our in-house scripts, all repeats formed a non-redundant list of the Chinese monal repeats. Microsatellites were detected using Krait (v0.10.2) [17] and the total length of perfect SSR (Simple Sequence Repeats) was calculated.

2.3.2. Non-coding RNAs annotation

Four types of non-coding RNAs (ncRNA) were determined by searching databases. Software, tRNAscan-SE (v1.3.1) [18], was used for reliable tRNA (transfer RNAs) positions. SnRNA (small nuclear RNAs) and miRNA (microRNAs) were detected through a two-step method covariance model of Rfam family. These were then aligned with BLAST and INFERNAL (v1.0.2) [19] to search for putative sequences in the Rfam database (Release 12.0). Red jungle fowl (*Gallus gallus*) full-length rRNA was used as queries for possible rRNA position detection in the Chinese monal genome within BLAST with an *E*-value cutoff of 1E-1.

2.3.3. Gene prediction

We used homology-based, *de novo* and transcriptome-based methods to predict the gene set in the Chinese monal genome. A *de novo* prediction was performed on the assembled genome with repetitive sequences masked as “N” based on the HMM (Hidden Markov Model) algorithm. After training the model on the protein sequences of the red jungle fowl genome, AUGUSTUS (v3.2.1) [20] and GENSCAN [21] programs were executed to find coding genes. For the homology prediction, proteins of the red jungle fowl, wild turkey (*Meleagris gallopavo*) and zebra finch (*Taeniopygia guttata*) from Ensembl 85 release were mapped onto the Chinese monal genome using TblastN (BLAST v2.2.28+) with an E-value cutoff of 1E-5. The results yielded from TblastN were joined by Solar (v0.9.6). Homologous sequences were successively aligned against the matching gene models using GeneWise (v2.4.1) [22]. 7.2 Gb RNA-seq data was also used to obtain the gene structures by PASA (v2.02) [23]. The three predictions above were integrated using EVM (v1.1.1) [24] to produce a consensus gene set. The integrated weights of the *ab initio*, protein, and transcript predictions were 10, 5, and 1 respectively. The three lines of results were manually inspected with Apollo (v1.11.6) [25] and the online Blastp to utilize a high quality final gene-set.

2.3.4. Functional annotation

The function of Chinese monal genes were assigned to annotated proteins based on the best match derived from the alignments in the SwissProt and TrEMBL (Uniprot release 2016.4.15) databases, using Blastp with a cutoff *E*-value of 10^{-5}. The motifs and domains of genes were annotated using InterProScan (v5.18-57.0) [26] against publicly available databases, including ProDom, PRINTS, PIRSF, Pfam, ProSiteProfiles, PANTHER, SUPERFAMILY and SMART. Descriptions of gene products from Gene Ontology ID were retrieved from the results of InterProScan. All genes were uploaded to KAAS (KEGG Automatic Annotation Server) to find the best match for each gene and the pathway in which the gene might be involved.

2.4. Genome evolution analysis

2.4.1. Gene family analysis

We used the TreeFam methodology [27] to define gene families from the eight bird species' genomes (Chinese monal, red jungle fowl, wild turkey, zebra finch, ground tit (*Pseudopodoces humilis*), mallard duck (*Anas platyrhynchos*), Japanese quail (*Coturnix japonica*), and ostrich (*Struthio camelus*)). Gene family expansion and contraction analysis was performed by CAFÉ (v1.5) [28] with α-priori *p*-value of < 0.05. KEGG annotation and enrichment analysis was performed by KOBAS (v3.0, http://kobas.cbi.pku.edu.cn/index.php).

2.4.2. Phylogenetic analysis

The phylogenetic relationship between the Chinese monal and seven chosen bird species genomes (red jungle fowl, wild turkey, zebra finch, ground tit, mallard duck, Japanese quail, ostrich) was reconstructed using single-copy orthologous gene families. Coding sequences (CDS) from each single-copy family were aligned by PRANK (v140603) and concatenated to one super gene for each species for building trees. “GTR + I + gamma” was selected as the best substitution model for nucleotide sequences in Modeltest. Then, RAxML (v8.2.8) [29] was applied to reconstruct ML phylogenetic trees with 1000 bootstrap replicates. Divergence time estimation analysis was performed by PAML MCMCTREE (v4.4) [30].

2.4.3. Positive selection analysis

The rate ratio (Ka/Ks, or ω) of nonsynonymous to synonymous nucleotide substitutions was estimated using PAML Codeml under the branch-site model to search the positively selected gene on aligned CDS of single copy orthologs in the Chinese monal genome. Potential positively selected genes were then filtered by FDR < 0.05. KEGG annotation and enrichment analysis was performed by KOBAS.

2.4.4. Temporal population dynamics of the Chinese monal

Effective population size of the Chinese monal was inferred by PSMC (v0.6.5). The mutation rate was calculated through the formula: $\mu = (\text{counts of mutated loci/sequence length})/2t$ (μ is mutation rate; t is generation time of the Chinese monal). Its substitution rate is 3.27e-09 substitutions per site per year. The generation time of Chinese monal is 4. PSMC's parameters were “-N25 -t15 -r5 -p ″4 + 25 * 2 + 4 + 6” with 100 bootstraps.

3. Results

3.1. Assembly and annotation of the genome

The final Chinese monal genome assembly was 1.01 Gb in length, which is close to the estimated genome size of 1.1 Gb (Table S1, Fig. S1). The genome covered by the genome assembly is 91.8%. The scaffold N50 and contig N50 of Chinese monal are 6.9 Mb and 12 Kb, respectively. The final Chinese monal genome assembly was compared with published bird species genomes and scaffold and contig N50 values were middle-range (Table 2). The CEGMA assessment showed that our assembly captured 82% (204) and 87% (215) complete and partial gene sets, respectively, of the 248 ultra-conserved Core Eukaryotic Genes (Table S2). BUSCO analysis indicated that 77% (329) of the 429 expected vertebrate genes were identified as complete single copy genes, while 20% (88) were considered missing during the assembly (Table S3). Missing genes in the Chinese monal (20%) were compared to the number of missing genes in the wild turkey (24%, 106) and consequently we argue that the monal's genome was sufficiently complete. The RNA sequencing mapping showed that 97.01% of the assembled transcripts were covered in the Chinese monal scaffolds. It is confirmed that the first assembly of the Chinese monal is of high quality and can be used as a reference genome for monal species.

The annotation showed a total length of 100.93 Mb (9.97%) as

ARTICLE IN PRESS

Table 2
Key genome assembly data of the Chinese monal compared to other sequenced birds.

Species	Total sequence length	Sequencing technology	Number of scaffolds	Scaffold N50 (bp)	Number of contigs	Contig N50	references
Chinese monal	1,012,210,055	NGS	2166	6,956,351	12,778	226,368	–
Turkey[a]	1,128,339,136	NGS	233,806	3,801,642	296,331	26,671	[31]
Red jungle fowl	1,230,258,557	NGS	23,870	6,379,610	24,693	2,894,815	[32]
Zebra finch	1,232,135,591	NGS	37,422	8,236,790	124,806	38,639	[33]
Ground tit	1,042,997,632	NGS	5406	16,337,386	27,052	165,265	[34]
Duck	1,105,052,351	NGS	78,488	1,233,631	227,449	26,114	[35]

[a] All the summary statistics of other birds from Enseml 85 release.

Table 3
Repetitive elements statistics in the Chinese monal and three other avian genomes.

Type	Chinese monal		Turkey		Red jungle fowl		Zebra finch	
	Length (bp)	Percentage(%)	Length (bp)	Percentage(%)	Length (bp)	Percentage(%)	Length (bp)	Percentage (%)
DNA[a]	10,133,591	1.0011	9,745,260	0.9176	11,493,786	1.0369	4,518,987	0.3665
LINEs	72,583,572	7.1708	79,690,942	7.5040	101,893,777	9.1922	71,838,312	5.8255
SINEs	3,490,274	0.3448	568,874	0.0536	927,195	0.0836	1,182,807	0.0959
LTR	12,539,167	1.2388	23,659,710	2.2279	48,207,392	4.3490	89,380,177	7.2480
Other	5887	0.0006	2310	0.0002	3483	0.0003	21,094	0.0017
Unclassified	2,174,157	0.2178	741,867	0.0699	1,944,347	0.1764	2,279,044	0.1848
Total	100,926,648	9.9710	98,108,450	9.2382	135,777,107	12.2488	148,766,379	12.0637

[a] DNA: DNA transposons; LINE: Long Interspersed Nuclear Elements; SINE: Short Interspersed Nuclear Elements; LTR: Long Terminal Repeated Elements.

being repetitive. The differences in repeat elements between other genomic sequenced birds are shown in Table 3 and Fig. S2. The total length of perfect SSR was found to be 5,465,549 bp (0.54%). Summaries of length and type are detailed in Table S4 and Table S5. We identified a total of 785 ncRNAs including 397 miRNAs, 15 rRNAs, 105 snRNAs, and 268 tRNAs (Table S6). The non-redundant Chinese monal protein-coding gene set contains about 16,940 genes (Table S7, Fig. S3), of which 15,550 genes demonstrated significant similarity matches to the public function databases of SwissProt, TrEMBL, GO, KEGG, and InterPro, of which 14,854 (87.7%), 15,466 (91.3%), 9303 (54.9%), 9679 (57.1%), 14,961 (88.3%), respectively, were matches (Table S8, Fig. S4).

3.2. Phylogenetic analysis

We found that the Chinese monal and wild turkey (Phasianidae) share a closer relationship than other Phasianids analyzed, with a divergence time of about 25.4 million years (Fig. 2). This is consistent with the complete mitochondrial genome of the Chinese monal [3] from our previous studies. In addition, there were more genomic structure (the types of repeat elements and average exon length) similarities between the Chinese monal and the wild turkey than the other six species analyzed. Orthologous genes in the Chinese monal and seven other bird species are shown in Table S9.

3.3. Temporal population dynamics

We compared Chinese monal effective population dynamics with the global climate history (PSMC analysis). The Chinese monal's population has expanded once and contracted once since the species' origin (Fig. S5). We found that global climatic change influenced a decline in the Chinese monal population in the middle of the last glaciation period. The population was almost halved from its original number 10,000 years ago, which is a considerable decline. A previous study of the population history of 36 bird species, using PSMC, found that other threatened birds experienced a similar decline during the last glacial period [36]. Thus, global climate change in last glaciation period exerts a strong influence on Chinese monal population numbers and other threatened birds with similar historic declines.

3.4. Evolution of expanded gene families and positive selection

Comparative genomic analysis with the other seven avian genomes demonstrated that the Chinese monal has a large number of genes involved in environmental adaptation. There were 297 species-specific genes (Table S10), 325 expanded gene families (Table S11) and 737 positive selected genes (Table S12) in the Chinese monal genome. We found that the species-specific genes, expanded gene families and positive selection genes in KEGG enriched pathways were mainly linked to growth and development, metabolism and hypoxia response that are associated with the high altitude adaptation of the Chinese monal.

3.4.1. Growth and development

The Chinese monal is the largest of the three monals and, by mass, is one of the largest pheasants (after the turkeys and the green and Indian peafowls). Results of the gene family and positive selection analysis indicated that 11 enrichment pathways were associated with the growth and development of the Chinese monal. Of these 11, some of the pathways have direct or indirect relationships as shown in Fig. 3. We found three important signaling pathways are involved in transduction and regulation; these were 1) MAPK signaling pathway (Fig. S6), 2) ErbB signaling pathway (Fig. S7) and 3) Wnt signaling pathway (Fig. S8). 1) The mitogen-activated protein kinase (MAPK) pathway is involved in various cellular functions, including cell proliferation, differentiation and migration [37]. We found CACN, GF and RTK genes in the classical MAPK pathway were positively selected. Similarly, PTP and MAFAPK were positively selected and are associated with the function of p38. 2) Within the ErbB signaling pathway, the ErbB family is a kind of typical cell membrane receptor and it contains four receptor tyrosine kinases, structurally related to the epidermal growth factor receptor (EGFR), involved in regulating cell proliferation, migration, differentiation, apoptosis and cell mobility [38]. ErbB 4, STAT5 and p27 are expanded, and AR and GSK-3 are positively selected in the Chinese monal. 3) The conservative Wnt/β - Catenin pathway regulates stem cell pluripotency and determines the cell differentiation in the development. The developmental cascade integrates other signal pathways, including retinoic acid, fibroblast growth factor (FGF), transforming growth factor beta (TGF - β) and BMP (BMP), and exists in a variety of different cell types and tissues [39, 40]. In this pathway, the wnt gene family has both selected and expanded genes and is a secreted

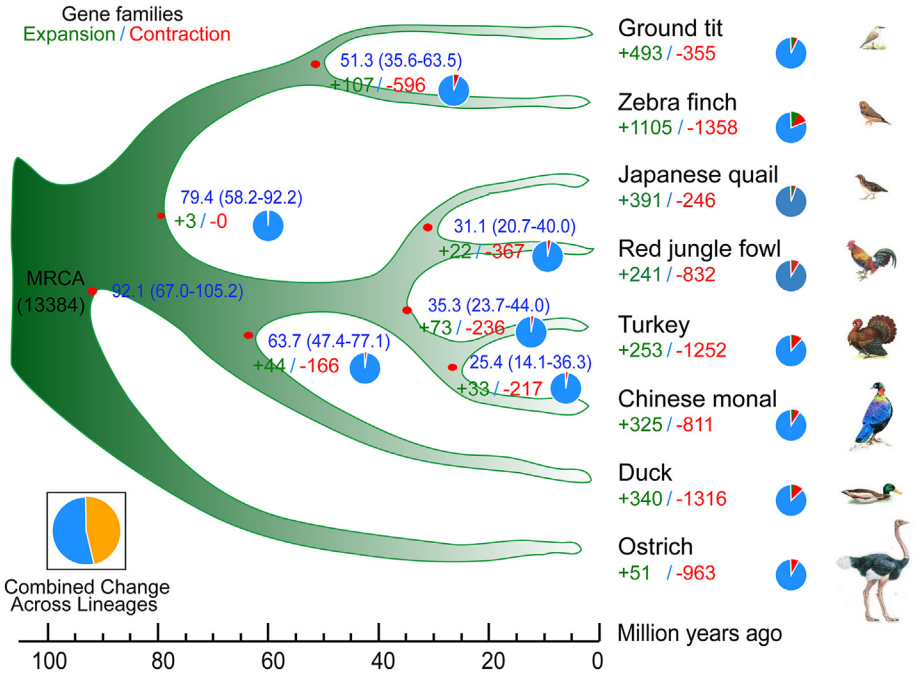

Fig. 2. Phylogenetic relationship and distance of the Chinese monal and seven other bird species.

glycoprotein. This wnt gene family can be combined with Frizzled (crimp) receptors, and then with LRP5/6 on the cell surface to form a complex to regulate downstream genes. The Frizzled family is also expanded and the three upstream genes (GSK-3β, APC, PS-1) of β-catenin and TBL1 are selected. In addition to these three important pathways, we found other enrichment pathways such as the ECM-receptor interaction (Fig. S9). This ECM-receptor interaction has positively selected, lineage-specific and gained genes, and works through a variety of signaling pathways in cell cycle regulation, adhesion, migration, and embryonic development. These expanded gene families and positive selection play important roles in many aspects of the development and growth of the Chinese monal.

3.4.2. Metabolism

High altitudes are very cold, thus adaptive thermoregulation is necessary for normal metabolism and daily activities. The Chinese monal has 12 enriched pathways of expanded gene families and positively selected genes associated with carbohydrate and lipid metabolism, and nucleotide and amino acid metabolism. Temperature regulation is a result of the process of metabolism. The three major metabolites (carbohydrate, lipid and amino acid) release energy in the process of oxidation, of which about 50% of the body's energy budget is used to maintain body temperature. Therefore, our findings of enrichment of xenobiotics metabolism by cytochrome P450, steroid hormone biosynthesis and retinol metabolism presented in expanded gene families are likely to be involved, and advantageous, in allowing the Chinese monal to live in the cold highlands.

3.4.3. Hypoxia response

The Chinese monal inhabits environments between 3500 and

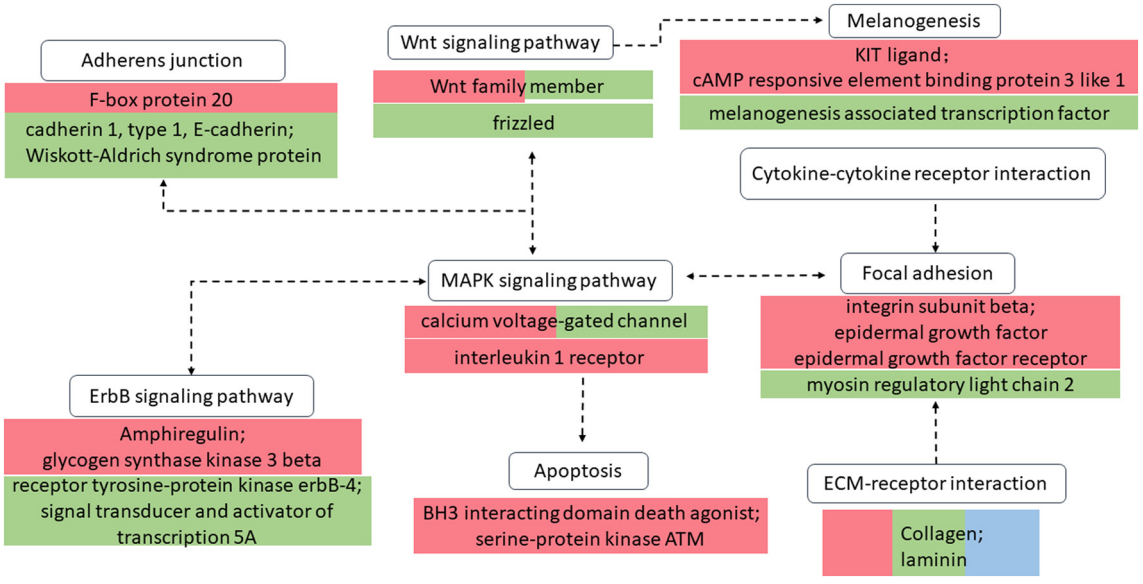

Fig. 3. Pathway relationships of expanded gene families and positively selected genes in Chinese monal.

4500 m above sea level (asl) of the Qinghai-Tibet Plateau. Generally, the saturation of oxyhemoglobin begins to decrease rapidly in the body above 2100 m asl due to lower atmospheric oxygen concentrations [41]. Animals need to maintain oxyhemoglobin saturation to ensure sufficient oxygen reaches bodily tissues, therefore it is highly advantageous for animals to have adaptations to avoid hypoxia and drops in oxyhemoglobin saturation above 2100 m asl. The Chinese monal lives in environments above this saturation threshold and thus adaptations to avoid hypoxia would be crucial for survival at these altitudes. We found 22 positively selected genes (Table S13) and one species-specific enriched gene family that are associated with hypoxia, presumably an adaptation that has (partly) allowed the Chinese monal to survive in a low oxygen environment. The HIF - 1 signaling pathway is an important hypoxic response pathway, where HIF-1-α acts as a master regulator of numerous hypoxia-inducible genes under hypoxic conditions [42]. The positively selected gene HIF1AN is the HIF-1-α inhibitor, which is also enhanced in the Tibetan people, Tibetan chicken, yak and ground tit. The HIF-target genes, VEGF family (vascular endothelial growth factor), ANGPTL4 and ANGPTL5 [43], APOD, EDNRA [44, 45] were also found under positive selection. These genes are involved in vasculogenesis (the formation of the circulatory system) and angiogenesis (the growth of blood vessels from pre-existing vasculature) [42, 46, 47]. These genes are advantageous as they increase O_2 delivery, mediate adaptive responses to O_2 deprivation and regulate a series of downstream target genes related to hypoxia response, including vasodilatation, nitric oxide synthesis, and cell respiration. The VEGF signaling pathway were also enriched in the lineage-specific genes [47, 48]. Of the 22 positive selected genes related to hypoxic adaptation, EDNRA, ADAM family [43, 49], DNAJ family and other genes associated with platelet production were also identified under positive selection.

3.4.4. Melanogenesis

The enrichment of expanded families in the melanogenesis pathway (Fig. S10) and Wnt signaling pathway are potentially beneficial to melanin and plumage color regulation in the Chinese monal [50, 51]. The gained genes in the MITF gene family, a downstream signaling gene of chromatophore, also plays a role in the development, differentiation and functional regulation of the pigment cell [52]. Two upstream genes, SCF and CREB, also had positive selection.

4. Discussion

4.1. Adaptation to the climatic conditions of the Qinghai-Tibet plateau

Low temperature, low oxygen and high UV are the three most significant climate characteristics of high-altitude environments of the Qinghai-Tibet plateau. The Chinese monal has inhabited this cold, harsh and high-altitude environment for millennia and therefore has acquired many adaptive life strategies for survival. We found many molecular adaptation mechanisms related to low temperatures, low oxygen and high UV to enhance the survival of the Chinese monal in its Qinghai-Tibet plateau environment.

We discovered that the Chinese monal has underlying molecular mechanism that has evolved in response to low temperatures. We found an increase of thermogenesis and heat preservation through enhanced metabolism and increased fat accumulation. The Chinese monal has enrichment in a series of metabolic activities, which were mainly related to carbohydrate, lipid and amino acid metabolism. Metabolic release of energy can ensure that the Chinese monal body is maintained at a relatively constant temperature and supported by other physiological and behavioral temperature regulation mechanisms. The underlying molecular mechanism has allowed the Chinese monal to grow larger and heavier, while maintaining a smaller skin surface area. This ratio of large mass to smaller surface area (*i.e.* Bergmann's rule) allows the Chinese monal to retain heat and expend less energy than an animal with the same mass but larger surface area. Retention of heat and more efficient use of energy is an effective survival strategy in the cold and sparse foraging potential of the Qinghai-Tibet plateau. This is very similar to the survival strategy of the ground tit, another iconic bird of the Qinghai-Tibet plateau.

Hypoxia adaptation in the Chinese monal is primarily related to the downstream regulation of blood vessels, hemoglobin and cell respiration through the HIF signaling pathway. The underlying HIF signaling pathway allows the Chinese monal to utilize lower oxygen concentrations and efficiently transfer oxygen around its body. Some researchers have suggested that birds have the innate ability of hypoxia adaptation in their breathing and oxygen transport systems, while other vertebrate oxygen transport systems are typically poorly adapted to low oxygen environments [53]. However, we found the same positively selected gene families in the Chinese monal (*e.g.* HIFAN, EDNRA) as in other birds and mammals, suggesting that at least some mammals may also possess the ability to inhabit low oxygen environments without significant adaptation [51]. A few well-known hypoxia adapted genes found in humans and other mammals (*i.e.* EPAS1, EGLN1, PTEN and PPARA) were not identified in the Chinese monal. These data can provide a good foundation for future comparative analyses of low oxygen adaptation in birds and mammals (Fig. S11).

Chinese monals have adapted to intense UV radiation *via* melanin-related metabolism and signal pathway expansion. We found that the Wnt pathway plays a crucial role in melanin regulation. The Wnt pathway is activated by changes in levels of UV radiation and this activation eventually leads to the adjustment of melanin content in the cell. Increased melanin protects cells from UV radiation damage. Similarly, the Wnt signaling pathway and melanin adjustment is associated with feather pigment metabolism in the Chinese monal.

Chinese monals have successfully adapted to the low temperature, low oxygen and high UV radiation of the Qinghai-Tibet plateau due to molecular mechanisms regulating carbohydrate, lipid and amino acid metabolism, efficient oxygen use and transport and effective adjustments to melanin content. This is the first study to identify these underlying molecular mechanisms and can provide a solid foundation, and reference material, for future research investigating bird and potentially mammal adaptations to high-altitude environments.

In conclusion, we have provided the first genome of the genus *Lophophorus*. The quality of *de novo* assembly and genomic annotation is good enough to be used as a reference genome for monal species. Various annotation resources that we have provided can be used to study comparative analyses between different species, especially in high-altitude adaptability, diversity of plumage color and molecular evolution of Galliformes species. Our genomic evolution analyses found considerable numbers of highland-adaptive genes that may help the Chinese monal survive in the harsh environment of the Hengduan Mountains. This research is preliminary and further work is required to fully explore the evolution of high-altitude adaptation, particularly comparisons of mammalian and avian adaptation to the Hengduan Mountains and broader Qinghai-Tibetan Plateau. In addition, considerable conservation effort is still required for the Chinese monal and that the results from this and future studies can assist in monitoring population health and conservation action success to recover populations.

Acknowledgements

This study was supported by grants from the National Natural Science Foundation of China (No. 31501851, 31772478), National Key Program of Research and Development, Ministry of Science and Technology (No. 2016YFC0503200) and the international visiting program for excellent young scholars of Sichuan University.

ARTICLE IN PRESS

Author's contributions

BY, YW, WL conceived the project. YW collected samples and extracted the genomic DNA.

KC, WL performed genome assembly. KC, WL, CY, JJ and CP performed the annotation. KC, CP performed phylogenetic analysis. BY, WL and CY, LD provided advice. KC, BY, YW, JG, ZF, MP lead the writing of the paper.

Data accessibility

Supporting data and materials are available in the NCBI bioproject accession number PRJNA321629.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2018.07.016.

References

[1] J. del Hoyo, A. Elliott, J. Sargatal, D.A. Christie, Handbook of the Birds of the World Alive, Lynx Edicions, http://www.hbw.com/node/53484, (2017) , Accessed date: 15 March 2017.

[2] M. Liu, Y. Gao, J. Chen, Z. Xiao, D.U. Chongyangzi, L.I. Minjie, J. Wang, L. Wang, W. Liu, C. Yang, Research status of conservation biology on *Lophophorus lhuysii*, Jiangxi Forestry Science and Technology, 2013.

[3] L.L. Ma, X.Y. Zhang, B.S. Yue, J.H. Ran, Complete mitochondrial genome of the Chinese Monal pheasant Lophophorus lhuysii, with phylogenetic implication in Phasianidae, Mitochondrial DNA 21 (2010) 5–7.

[4] F. He, Ecology of the Chinese monal in winter, Zool. Res. 6 (1985) 345–352.

[5] L.U. Taichun, Ecological studies on chinese monal, Acta Zool. Sin. 3 (1986) 011.

[6] BirdInternational, The IUCN Red List of Threatened Species, (2016), https://doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22679192A92806697.en Accessed 15 March 2017.

[7] M.J.B. Simmons, M.R.I. Beyer, Appendices I, II and III, https://cites.org/eng/app/appendices.php, (2017) , Accessed date: 2 January 2017.

[8] X. Zhan, Z. Zhang, A. Wu, Y. Tao, Phylogenetic relationships of monal pheasants Lophophorus inferred from sequences of mitochondrial cytochrome b gene, Zool. Res. 24 (2003) 337–342.

[9] G.J. Dyke, B.E. Gulas, T.M. Crowe, Suprageneric relationships of galliform birds (Aves, Galliformes): a cladistic analysis of morphological characters, Zool. J. Linn. Soc. Lond. 137 (2003) 227–244.

[10] L.R. Zuo, The Monals of southwestern China, China Scenic 5 (2015).

[11] B. Liu, Y. Shi, J. Yuan, X. Hu, H. Zhang, N. Li, Z. Li, Y. Chen, D. Mu, W. Fan, Estimation of genomic characteristics by analyzing k-mer frequency in de novo genome projects, Quant. Biol. 35 (2013) 62–67.

[12] J. Butler, I. Maccallum, M. Kleber, I.A. Shlyakhter, M.K. Belmonte, E.S. Lander, C. Nusbaum, D.B. Jaffe, ALLPATHS: de novo assembly of whole-genome shotgun microreads, Genome Res. 18 (2008) 810–820.

[13] B.J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P.D. Blood, J. Bowden, M.B. Couger, D. Eccles, B. Li, M. Lieber, De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity, Nat. Protoc. 8 (2013) 1494–1512.

[14] G. Parra, K. Bradnam, I. Korf, CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes, Bioinformatics 23 (2007) 1061–1067.

[15] F.A. Simão, R.M. Waterhouse, P. Ioannidis, E.V. Kriventseva, E.M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs, Bioinformatics 31 (2015) 3210–3212.

[16] N.S. Chen, Using RepeatMasker to identify repetitive elements in genomic sequences, Curr. Protoc. Bioinformatics 5 (2004) 4–10.

[17] L. Du, C. Zhang, Q. Liu, X. Zhang, B. Yue, Krait: an ultrafast tool for genome-wide survey of microsatellites and primer design, Bioinformatics 34 (2017) 681–683.

[18] T.M. Lowe, S.R. Eddy, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence, Nucleic Acids Res. 25 (1997) 955–964.

[19] E.P. Nawrocki, D.L. Kolbe, S.R. Eddy, Infernal 1.0: inference of RNA alignments, Bioinformatics 25 (2009) 1335–1337.

[20] M. Stanke, O. Keller, I. Gunduz, A. Hayes, S. Waack, B. Morgenstern, AUGUSTUS: ab initio prediction of alternative transcripts, Nucleic Acids Res. 34 (2006) 435–439.

[21] C. Burge, S. Karlin, Prediction of complete gene structures in human genomic DNA, J. Mol. Biol. 268 (1997) 78–94.

[22] E. Birney, M. Clamp, R. Durbin, GeneWise and Genomewise, Genome Res. 14 (2004) 988–995.

[23] B.J. Haas, A.L. Delcher, S.M. Mount, J.R. Wortman, R.K.S. Jr, L.I. Hannick, R. Maiti, C.M. Ronning, D.B. Rusch, C.D. Town, Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies, Nucleic Acids Res. 31 (2003) 5654–5666.

[24] B.J. Haas, S.L. Salzberg, W. Zhu, M. Pertea, J.E. Allen, J. Orvis, O. White, C.R. Buell, J.R. Wortman, Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments, Genome Biol. 9 (2008) 7.

[25] S.E. Lewis, S. Searle, N. Harris, M. Gibson, V. Iyer, J. Richter, C. Wiel, L. Bayraktaroglu, E. Birney, M.A. Crosby, Apollo: a sequence annotation editor, Genome Biol. 3 (2002) (research0082-1).

[26] A. Bairoch, R. Apweiler, C.H. Wu, W.C. Barker, B. Boeckmann, S. Ferro, The universal protein resource (UniProt), Nucleic Acids Res. 33 (2005) 154–159.

[27] H. Li, A. Coghlan, J. Ruan, L.J. Coin, J.K. Hériché, L. Osmotherly, R. Li, T. Liu, Z. Zhang, L. Bolund, TreeFam: a curated database of phylogenetic trees of animal gene families, Nucleic Acids Res. 34 (2006) 572–580.

[28] T. De Bie, N. Cristianini, J.P. Demuth, M.W. Hahn, CAFE: A computational tool for the study of gene family evolution, Bioinformatics 22 (2006) 1269–1271.

[29] A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies, Bioinformatics 30 (2014) 1312–1313.

[30] Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood, Mol. Biol. Evol. 24 (2007) 1586–1591.

[31] R.A. Dalloul, J.A. Long, A.V. Zimin, L. Aslam, K. Beal, B.L. Ann, P. Bouffard, D.W. Burt, O. Crasta, R.P. Crooijmans, Multi-platform next-generation sequencing of the domestic turkey (Meleagris gallopavo): genome assembly and analysis, PLoS Biol. 8 (2010) 1000475.

[32] L.D.W. Hillier, W. Miller, E. Birney, W. Warren, R.C. Hardison, C.P. Ponting, P. Bork, D.W. Burt, M.A.M. Groenen, M.E. Delany, Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution, Nature 432 (2004) 695–716.

[33] W.C. Warren, D.F. Clayton, H. Ellegren, A.P. Arnold, L.D.W. Hillier, A. Künstner, S. Searle, S. White, A.J. Vilella, S. Fairley, The genome of a songbird, Nature 464 (2010) 757.

[34] Q. Cai, X. Qian, Y. Lang, Y. Luo, J. Xu, S. Pan, Y. Hui, C. Gou, Y. Cai, M. Hao, Genome sequence of ground tit Pseudopodoces humilis and its adaptation to high altitude, Genome Biol. 14 (2013) R29.

[35] Y. Huang, Y. Li, D.W. Burt, H. Chen, Y. Zhang, W. Qian, H. Kim, S. Gan, Y. Zhao, J. Li, The duck genome and transcriptome provide insight into an avian influenza virus reservoir species, Nat. Genet. 45 (2013) 776–783.

[36] K. Nadachowskabrzyska, C. Li, L. Smeds, G. Zhang, H. Ellegren, Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences, Curr. Biol. 25 (2015) 1375–1380.

[37] R. J, MAP Kinase Pathway, Springer International Publishing, 2015 Textbook of Cell Signalling in Cancer.

[38] M.A. Olayioye, R.M. Neve, H.A. Lane, N.E. Hynes, The ErbB signaling network: receptor heterodimerization in development and cancer, EMBO J. 19 (2000) 3159–3167.

[39] H. Clevers, Wnt/β-catenin signaling in development and disease, Cell 127 (2006) 469–480.

[40] K.M. Cadigan, R. Nusse, Wnt signaling: a common theme in animal development, Genes Dev. 11 (1997) 3286–3305.

[41] A.J. Young, J.T. Reeves, Human adaptation to high terrestrial altitude, Med. Asp. Harsh Environ. 2 (2002) 647–691.

[42] G.L. Semenza, HIF-1 and mechanisms of hypoxia sensing, Curr. Opin. Cell Biol. 13 (2001) 167–171.

[43] Y. Qu, E. Al, Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan plateau, Sci. Found. China 4 (2014) 2071.

[44] Q. Qiu, G. Zhang, T. Ma, W. Qian, J. Wang, Z. Ye, C. Cao, Q. Hu, J. Kim, D.M. Larkin, The yak genome and adaptation to life at high altitude, Nat. Genet. 44 (2012) 946.

[45] W. Zhang, Z. Fan, E. Han, R. Hou, L. Zhang, M. Galaverni, J. Huang, H. Liu, P. Silva, P. Li, Hypoxia adaptations in the grey wolf (canis lupus chanco) from Qinghai-Tibet Plateau, PLoS Genet. 10 (2014) e1004466.

[46] F.J. Giordano, H.P. Gerber, S.P. Williams, N. Vanbruggen, S. Bunting, P. Ruizlozano, Y. Gu, A.K. Nath, Y. Huang, R. Hickey, A cardiac myocyte vascular endothelial growth factor paracrine pathway is required to maintain cardiac function, P. Natl. Acad. Sci. USA 98 (2001) 5780–5785.

[47] Y. Yamazaki, T. Morita, Molecular and functional diversity of vascular endothelial growth factors, Mol. Divers. 10 (2006) 515.

[48] N. Ferrara, H.P. Gerber, J. Lecouter, The biology of VEGF and its receptors, Nat. Med. 9 (2003) 669–676.

[49] T.S. Simonson, Y. Yang, C.D. Huff, H. Yun, G. Qin, D.J. Witherspoon, Z. Bai, F.R. Lorenzo, J. Xing, L.B. Jorde, Genetic evidence for high-altitude adaptation in Tibet, Science 329 (2010) 72–75.

[50] C. Duval, M. Régnier, R. Schmidt, Distinct Melanogenic Response of Human Melanocytes in Mono-culture, in Co-Culture with Keratinocytes and in Reconstructed Epidermis, to UV Exposure, Pigment Cell Res. 14 (2001) 348–355.

[51] S. Li, C. Wang, W. Yu, S. Zhao, Y. Gong, Identification of genes related to white and black plumage formation by RNA-Seq from white and black feather bulbs in ducks, PLoS ONE 7 (2012) e36592.

[52] C. Levy, M. Khaled, D.E. Fisher, MITF: master regulator of melanocyte development and melanoma oncogene, Trends Mol. Med. 12 (2006) 406–414.

[53] G.R. Scott, J.U. Meir, L.A. Hawkes, P.B. Frappell, W.K. Milsom, Point: high altitude is for the birds!, J. Appl. Physiol. 111 (2011) (2011) 1514–1519.